View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Apollo

Functional interactions between Polypyrimidine Tract Binding Protein and PRI peptide ligand containing proteins

[1]Miguel B. Coelho, [1]David B. Ascher, [1]Clare Gooding, [1]Emma Lang, [1]Hannah Maude, [1]David Turner, [1]Miriam Llorian, [1,3]Douglas E.V. Pires, [2]Jan Attig & [1]Christopher WJ Smith

[1]Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW
[2]Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, WC1N 3BG, UK
[3]Centro de Pesquisas René Rachou, Fundação Oswaldo Cruz, Avenida Augusto de Lima 1715, Belo Horizonte, 30190-002, Brazil

Tel: 01223-333655

Fax: 01223-766002

Email: cwjs1@cam.ac.uk

Keywords: alternative_splicing, Polypyrimidine_tract_binding_protein, PTB, Matrin3, hnRNP, RNA

Abstract

Polypyrimidine Tract Binding protein (PTBP1) is a heterogeneous nuclear ribonucleoprotein (hnRNP) that plays roles in most stages of the life-cycle of pre-mRNA and mRNAs in the nucleus and cytoplasm. PTBP1 has four RNA binding domains of the RNA Recognition Motif (RRM) family, each of which can bind to pyrimidine motifs. In addition, RRM2 can interact via its dorsal surface with proteins containing short peptide ligands known as PTB RRM2 Interacting (PRI) motifs, originally found in the protein Raver1. Here we review our recent progress in understanding the interactions of PTB with RNA and with various proteins containing PRI ligands.

Introduction

PTBP1 acts widely as a regulator of alternative splicing, primarily as a repressor but also as an activator depending upon where it binds relative to a regulated exon [1, 2]. It also regulates use of some 3´ end processing/polyadenylation sites, internal ribosome entry site (IRES) driven translation, cytoplasmic localization and mRNA stability (reviewed in [3]), including by influencing Nonsense Mediated mRNA Decay (NMD) [4]. Strikingly, knockdown of PTBP1 promotes conversion of fibroblasts to neuron-like cells, principally by affecting a network of miRNA-regulated mRNAs [5]. In addition to PTBP1, mammals possess two paralogs — PTBP2 (also known as nPTB or brPTB) and PTBP3 (also known as ROD1), both of which are regulated post-transcriptionally by PTBP1 (Fig. 1). Skipping of PTBP2 exon 10 is promoted by PTBP1, leading to frameshifting and NMD, thereby restricting PTBP2 expression to cells where PTBP1 levels are reduced (reviewed in [6-8]). PTBP1 and PTBP2 also promote exon 2 skipping in PTBP3, leading to production of N-terminally truncated protein isoforms initiating at internal AUG codons 4 and 11. Both these isoforms show a higher degree of cytoplasmic localization, due to loss of the bipartite nuclear localization signal, and the shorter AUG11 isoform also has a truncated RRM1 domain [9, 10](Fig.1). While the switch from PTBP1 to PTBP2

expression is known to play an important role in neuronal differentiation [8, 11, 12], the key physiological roles of PTBP3 and of the PTBP1/2 regulation of PTBP3 isoforms are not yet known, although from its expression patterns PTBP3 is thought to have roles in hematopoietic cells.

There are many open questions about the molecular and physiological functions of the PTBP proteins. These include the extent to which the PTBP1, 2 and 3 paralogs and their isoforms have unique non-redundant activities which are important in various developmental settings. While much progress has been made in understanding the PTBP1/2 axis in neuronal biology (reviewed in [8]), very little is known about PTBP3, which may play equally important roles in hematopoietic cells. Another key set of questions concern the molecular basis of PTBP1 regulation of various post-transcriptional processes. Although interaction with target pre-mRNAs/mRNAs may in some cases be sufficient for activity, by specifically blocking access of other factors, in other cases PTBP1 also interacts with other proteins or snRNAs to effect its function. Here we focus on PTBP1 activity as a splicing regulator.

RNA binding by PTB proteins

The PTBP proteins possess four RRM domains, with an N-terminal extension containing nuclear localization and nuclear export sequence motifs, flexible linkers between RRMs 1, 2 and 3 (Fig. 2A), and a linker between RRMs 3 and 4 that contributes to back-to-back packing of the two domains [13]. All four RRMs are able to bind RNA, with recognition of a core CU motif by RRMs 1, 2 and 3 [13, 14]. Crosslinking immunoprecipitation (CLIP) studies have shown widespread binding of PTBP1 to RNA, particularly in introns [2, 15], and integration with transcriptome-wide data for PTBP1 knockdown has shown how the position of binding on the pre-mRNA relates to its activity as a repressor or activator. PTBP1 binding upstream of and within regulated exons leads to repression, while binding downstream of the exon is associated with activation of splicing [1, 15].

Complementing the transcriptome-wide analyses we carried out a detailed analysis of how the individual RRMs of PTBP1 dock onto a well characterized target RNA: the PTBP1 repressed exon 6 of human *FAS* pre-

mRNA [16]. We used a combination of mutations designed to impair RNA binding by individual RRMs [17], along with a panel of mutants each of which has a single cysteine residue strategically located adjacent to the RNA binding surface of one of the RRMs, allowing mapping of the contact sites by tethered hydroxyl radical probing [18]. We found a distinctly uneven division of labour, with the C-terminal RRMs 3 and 4 contributing more to regulation of *FAS* exon 6 than RRMs 1 and 2 [19]. Surprisingly, RRM4, which has the lowest intrinsic specificity recognising only a YC dinucleotide [13], is the most important RRM for PTBP1 function, consistent with an earlier report [20]. Mutation of YC contact points on the RNA prevented RRM4 binding at that site but did not impair function, presumably due to the ready availability of alternative nearby YC sites. In contrast, mutations of RRM4 that impair RNA binding affect splicing repressor activity more severely than mutation of any other RRM [19]. RNA binding by RRM3, di-domain packing of RRMs 3 and 4, and formation of a surface electropositive patch upon didomain packing [21], which extends the binding surface for RNA bound at RRM4, were also found to be important for function [19]. RRMs 1 and 2 showed fewer and weaker binding contacts with *FAS* pre-mRNA. Sharma and colleagues have found that RRMs 1 and 2 contact stem-loop IV of U1 snRNA bound at the 5´ splice site of the *CSRC* N1 exon [22]. If this is a general mechanism of repression it could explain the lower level of contact of these RRMs with *FAS* pre-mRNA.

Protein binding by PTBP1

While much attention has understandably focused upon its interactions with RNA, PTBP1 is also involved in a number of protein-protein interactions (for systematic analyses, see [23, 24]). Some of the reported interactions might be RNA-mediated, possibly reflecting indirect interactions bridged by common RNA targets. Co-association on the same RNA by PTBP1 and other proteins could be functionally important even in the absence of direct protein-protein interactions. For example a cytoplasmic RNA-bound complex of PTBP1 with YBX1, PSF and NONO/p54 promotes IRES driven translation of a subset of mRNAs in response to apoptotic stimuli [25]. Moreover, we found that PTBP1 interacts directly with muscleblind like protein 1 (MBNL1) in an RNA-

dependent manner, requiring RNA binding only by MBNL1, suggesting that RNA binding induces a conformational change in MBNL1 to promote interaction with PTBP1, which could be important in a subset of co-regulated alternative splicing events (ASEs) [26]. For the majority of reported PTBP1-protein interactions the details of the interaction remain sketchy. An exception is the interaction of RRM2 with short peptide ligands.

Peptide mediated interactions with PTBP1 RRM2

Our focus upon RRM2 arose from structure-function analysis using an artificial MS2 tethering assay [27]. A PTBP1 binding site downstream of *Tpm1* exon 3 was replaced by a bacteriophage MS2 coat protein binding site, leading to loss of PTBP1 mediated repression. Exon skipping was restored by expression of PTBP1-MS2 fusion proteins, including a deletion mutant retaining just RRM2 and the following interdomain linker (referred to as RRM2L). Further deletion of either the linker or RRM2 abolished activity. The linker contains the site of a 26 amino acid insert that differentiates the major expressed PTBP1 isoforms from each other (Fig. 1, 2A), and the longer isoform of RRM2L showed stronger activity in the MS2 assay [27]. Recently, the enhanced activity of the longer isoform was shown by the Blencowe lab to apply to numerous alternative splicing events (ASEs), and was particularly relevant for a set of ASEs that change early during neuronal differentiation, in response to the shift from the long to the short linker isoform of PTBP1 [28]. The generation of long and short PTBP1 isoforms with differential activity was shown to be an evolutionary innovation acquired by mammals but not other vertebrate lineages [28]. Interestingly, PTBP2 lacks the exon responsible for the longer variant of the linker, while in PTBP3 this segment is constitutively included. Although this is not the only site of variation it is an interesting possibility that this region might influence differential activities between PTB paralogs.

RRM2 was first found to interact with proteins as well as RNA by investigation of the PTB-interacting protein Raver1 [29], overexpression of which promotes Tpm1 exon 3 skipping [30]. Interaction of Raver1 with PTB is mediated by 4-5 short heptapeptide motifs — termed PTB RRM2 Interaction (PRI) motifs. The Raver1 PRI motifs are sufficient for PTB interaction and

necessary for Raver1 activity [31]. Raver1 PRI1 (GLLGAPP) and PRI3 (SLLGEPP) bind PTBP1 with higher affinity than PRI2 and 4, and are conserved in the paralog Raver2 [32]. Limited mutagenesis of PRI3 established a consensus PRI motif of [S/G][I/L]LGx ϕP, where ϕ is a small hydrophobic side-chain [31, 33]. NMR analysis established that the PRI3 peptide interacted with RRM2 on the opposite face from the RNA binding surface, and that RRM2 could form a ternary complex with RNA and PRI peptide [31]. Crystallographic analysis of PRIs covalently linked by a flexible tether to RRM2 provided high-resolution structures of RRM2 bound to Raver1 PRI3 (1.4 Å) and PRI4 (1.55 Å) [33]. The PRI peptide interacts with the alpha helical face of RRM2 perpendicular to the two helices, similar to the mode of interaction of U2AF homology motif (UHM) domains of $U2AF_{65}$ and $U2AF_{35}$ with their UHM ligand motif (ULM) peptides [34, 35]. While ULMs have a characteristic tryptophan residue that docks into a pocket on the UHM dorsal surface, the PRI di-leucine motif docks into a shallower hydrophobic pocket on RRM2. PTBP1 Tyrosine 247 is a particular focus of hydrophobic and carbon-pi interactions and the di-Proline motif of PRI3 fits in a pocket between Tyr247 and Tyr193 (Fig. 2A). Mutation of tyrosine 247 to glutamine (Y247Q) impaired interaction of RRM2 with PRI peptides but not with RNA, confirming the importance of this residue to protein-protein interactions. Moreover, the mutation reduced the activity of RRM2L-MS2 in the tethered function assay arguing for the functional importance of PRI-binding by PTBP1 RRM2 [33].

Despite the detailed structural analysis of Raver1-PTBP1 interactions, a Raver1 knockout mouse showed no obvious alternative splicing defects [36]. Moreover, knockdown of HeLa cell Raver1 had only modest effects upon the transcriptome, with no observed upregulation of Raver2 which might have compensated for loss of Raver1 (M. Hallegger & CWJS, unpublished observations). Nevertheless, the specificity of the PTBP1-PRI interaction and the conservation of the PRI binding surface in the PTBP2 and PTBP3 paralogs [33] (Fig. 3C), suggested that it mediates functional interactions. We therefore carried out pull-down experiments from HeLa cell nuclear extracts using wild type (WT) and Y247Q PTBP1 RRM2 [15]. Numerous proteins in the nuclear extract interacted with WT but not Y247Q RRM2, the two most

prominent of which were Raver1, as expected, and the nuclear matrix protein Matrin3. Several of the interacting proteins contained sequences that match the established PRI consensus (Fig. 2B), and that interacted with GST-PTB (Fig 2C). These proteins were associated with various nuclear functions, with some of which PTBP1 has previously been associated. For example, PTBP1 regulates some 3´ end processing sites [37], and two 3' end processing factors, WDR33 and CSTF2, with PRIs were identified in the pull-down (Fig 2B). This suggests that PTB could activate selected polyA sites by binding nearby and helping to recruit WDR33 or CSTF2. Two other interesting PRI motif containing binding partners were the paralogs CCAR2/DBC1 and CCAR1 [38], both of which, among numerous other functions, have been associated with splicing. CCAR2 can help RNA polymerase II overcome difficult to transcribe regions and thereby alter exon inclusion co-transcriptionally [39], while CCAR1 is a spliceosomal A-complex component [23]. They share several domains, and in the N-terminus both have at least one PRI motif (Fig 2B,C). Both proteins represent potentially interesting targets for PTBP1's roles in splicing regulation.

Matrin 3 – a novel regulator of alternative splicing

The most abundant RRM2 interacting protein we identified was Matrin3 [15], which was originally identified as a prominent nuclear matrix protein (reviewed in [40]). It has a pair of DNA binding binding C_2H_2 Zn finger domains and two RNA binding RRM domains. Matrin3 has been suggested to play roles in a number of nuclear processes, but clear evidence in support of a direct role has been lacking in most cases. Given its association with a well-known splicing regulator we tested whether Matrin3 plays a role in regulating AS. By transcriptome analysis of Matrin3 knockdown in HeLa cells we found that Matrin3 acted widely as a splicing regulator, principally as a repressor of cassette exons [15]. When compared with previous data on PTB target exons [1], only 18% of Matrin3 regulated events were also targets of PTB, showing that for the majority of splicing events Matrin3 acts independently of PTB. In retrospect, this is perhaps not surprising; RBM20, which is one of Matrin3's closest relatives [40], has been found to regulate AS in cardiomyocytes [41,

42] even though it has no reported interactions with PTBP1 and does not possess a PRI motif. This supports the concept that Matrin3 and related proteins can operate as independent RNA-binding splicing regulators. Integration of Matrin3 iCLIP data with the transcriptome data revealed that Matrin3 is a direct regulator of exons that it represses and that it binds within the exons and within extended regions (up to 500 nt) of the flanking introns, which are larger on average than the introns flanking other cassette exons. Optimal binding motifs for Matrin3's RRM domains were enriched in the adjacent 250 nt of the introns to each side, but not within the cassette exons themselves. Finally, we found that Matrin3's RRM domains and PRI were necessary for regulation of two model ASEs, but the DNA binding Zn fingers were dispensable. Taken together with the known self-association of Matrin3 [43], the data suggest that Matrin3 acts by binding at high affinity sites followed by spreading to create a wide silenced region of RNA. Such models have been suggested previously for other splicing regulators (e.g. [44]), but Matrin3 is the first for which a splicing map derived from transcriptome-wide data suggests that this is a general mechanism. We are currently testing this model for Matrin3 action, as well as addressing why Matrin3 needs its PRI motif to regulate ASEs that are independent of PTBP proteins.

A wider network of PRI-RRM interactions?

There are suggestions that both sides of the PTB-PRI interaction may be more diverse than we currently know. On the RRM side stands the intriguing observation that Matrin3 requires its PRI motif, even for ASEs that are not co-regulated by PTBP1, suggesting that this PRI can interact with proteins other than PTBP1. On the PRI side, we know that many proteins interact with PTBP1 RRM2 in a Y247Q sensitive way, despite lacking a recognisable PRI motif. One explanation is that some of these proteins could be piggy-backing on direct interactors that possess a PRI motif. A second possibility is that some direct interactors may contain PRI motifs that do not conform to the consensus established by limited mutagenesis [31, 33].

As a first step to addressing the potentially wider spectrum of RRM-PRI interactions we have used a structural bioinformatics approach. The mCSM web-server was designed to predict the effects of missense variants upon

protein domain stability, but has been extended to predict effects upon protein-protein (mCSM-PPI), protein-nucleic acid (mCSM-NA) and protein-ligand (mCSM-Lig) interactions [45-47]. Using the structure of RRM2 bound to Raver1 PRI3 [31, 33], computational saturation mutagenesis was carried out to predict the effects upon RRM-PRI interaction of mutations to all other amino acids of each residue of the PRI motif and of the PRI interaction surface of RRM2. From the PRI side the mCSM-PPI predictions agreed with alanine scanning mutagenesis data from V498-P505 (Fig 3A) [31, 33]. Of particular interest, a number of previously untested sequence variants are predicted to increase the affinity of interaction, including aromatic side-chains in place of the di-leucine motif and also, more surprisingly, replacement of Proline 505 by aspartate or glutamate. Encouragingly, some of the Y247Q sensitive RRM2 interacting proteins contain potential PRI motif variants suggested by mCSM-PPI predictions. For example SUGP2, which was identified by GST-PTBP1 pull-downs from HeLa nuclear extracts [15] (Fig. 2B), in co-immunoprecipitation from HEK 293T cells (MC & CWJS unpublished observations), and in a global interactome survey [24], has the motif SLLGKGE. We are currently testing the functionality of some of these predicted novel PRI motif variants.

mCSM-PPI was also used to explore the effects of sequence variants in the 12 positions that constitute the peptide binding surface of PTBP1 RRM2. The critical role of Y247 was highlighted by the predicted deleterious effects of all mutations at this site (Fig 3B). Likewise, all mutations at Y193, V199, L241, N245, I246, N248 and L253 are predicted to impair PRI interaction. The PRI interacting residues are fully conserved between PTBP1, 2 and 3, and only one position (N248D) varies in the rodent restricted smPTB, [48]. Moreover, comparison with the RRM2 domains of hnRNPL and hnRNPLL, which are the next closest relatives (44 and 41% identity) and share characteristic structural features [49], shows that 6-7 of the positions are conserved, all of these corresponding to positions at which all mutations in PTBP1 would impair PTB-PRI interaction (Fig 3B,C). Moreover, in 4 of the 6 positions in hnRNPL/LL that vary from PTBP1, some mutations in PTBP1 are predicted to improve interaction. This suggests that the hnRNPL family proteins might also interact with co-regulators via their RRM2 domains, and

that these interacting proteins will also possess short linear peptides similar to the PRI motifs. Given the similarity, but non-identity, of their predicted peptide-binding surfaces, it might be expected that some interacting proteins might be shared between PTB and hnRNPL proteins, while others might be unique to each protein family. Indeed, Matrin3 is known to interact with hnRNPL [43, 50] although there is no evidence that this is mediated by the Matrin3 PRI. Thus, rather than being an interaction module unique to PTB, the PTB-PRI interface might be a prototype for a more extended network of protein-protein interactions, perhaps similar to the set of interactions between the UHM domains of U2AF1, U2AF2, SPF45, PUF60, KIS and HCC1 and the ULM ligands of SF3B1, SF1 and U2AF2 [34]. If this is the case, it raises the question of how the appropriate RRM-PRI interactions are selected. A clue to the answer is provided by the low affinity ($K_D \sim 100\ \mu M$) of the interaction between PTBP1 RRM2 and the strong PRI3 from Raver1 [31]. This suggests that PRI-RRM interactions might function in the context of multivalent interactions in multi-component regulatory RNP assemblies. We are addressing these possibilities in our continuing investigation of the extent and roles of PRI-RRM interactions.

Acknowledgements

Work in the CWJS lab was funded by grants from the BBSRC (BB/H004203/1), and from the Wellcome Trust (092900). JA was funded by a fellowship of the Boehringer Ingelheim Fond. DBA and DEVP are funded by The Medical Research Council (MR/M026302/1) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais. DBA is supported by an NHMRC CJ Martin Fellowship (APP1072476). We thank Jernej Ule for invaluable advice and input to our work.

Abbreviations

AS	Alternative splicing
ASE	Alternative splicing event
CLIP	Crosslinking immunoprecipitation

hnRNP	heterogeneous nuclear ribonucleoprotein
IRES	internal ribosome entry site
MBNL	Muscleblind like
NMD	Nonsense Mediated mRNA Decay
PRI	PTB RRM2 Interacting
PTBP1	Polypyrimidine Tract Binding Protein 1
PTBP2	Polypyrimidine Tract Binding Protein 2
PTBP3	Polypyrimidine Tract Binding Protein 3
RRM	RNA Recognition Motif
UHM	U2AF Homology Domain
ULM	UHM ligand motif

References

1 Llorian, M., Schwartz, S., Clark, T. A., Hollander, D., Tan, L.-Y., Spellman, R., Gordon, A., Schweitzer, A. C., de la Grange, P., Ast, G. and Smith, C. W. J. (2010) Position-dependent alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB. Nature structural & molecular biology. **17**, 1114-1123

2 Xue, Y., Zhou, Y., Wu, T., Zhu, T., Ji, X., Kwon, Y. S., Zhang, C., Yeo, G., Black, D. L., Sun, H., Fu, X. D. and Zhang, Y. (2009) Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. Mol Cell. **36**, 996-1006

3 Sawicka, K., Bushell, M., Spriggs, K. A. and Willis, A. E. (2008) Polypyrimidine-tract-binding protein: a multifunctional RNA-binding protein. Biochem Soc Trans. **36**, 641-647

4 Ge, Z., Quek, B. L., Beemon, K. L. and Hogg, J. R. (2016) Polypyrimidine tract binding protein 1 protects mRNAs from recognition by the nonsense-mediated mRNA decay pathway. eLife. **5**

5 Xue, Y., Ouyang, K., Huang, J., Zhou, Y., Ouyang, H., Li, H., Wang, G., Wu, Q., Wei, C., Bi, Y., Jiang, L., Cai, Z., Sun, H., Zhang, K., Zhang, Y., Chen, J. and Fu, X. D. (2013) Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated microRNA circuits. Cell. **152**, 82-96

6 McGlincy, N. J. and Smith, C. W. (2008) Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? Trends Biochem Sci. **33**, 385-393

7 Makeyev, E. V. and Maniatis, T. (2008) Multilevel regulation of gene expression by microRNAs. Science. **319**, 1789-1790

8 Keppetipola, N., Sharma, S., Li, Q. and Black, D. L. (2012) Neuronal regulation of pre-mRNA splicing by polypyrimidine tract binding proteins, PTBP1 and PTBP2. Critical reviews in biochemistry and molecular biology. **47**, 360-378

9 Tan, L. Y., Whitfield, P., Llorian, M., Monzon-Casanova, E., Diaz-Munoz, M. D., Turner, M. and Smith, C. W. (2015) Generation of functionally distinct isoforms of PTBP3 by alternative splicing and translation initiation. Nucleic acids research. **43**, 5586-5600
10 Spellman, R., Llorian, M. and Smith, C. W. (2007) Crossregulation and functional redundancy between the splicing regulator PTB and its paralogs nPTB and ROD1. Mol Cell. **27**, 420-434
11 Makeyev, E. V., Zhang, J., Carrasco, M. A. and Maniatis, T. (2007) The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. Mol Cell. **27**, 435-448
12 Boutz, P. L., Stoilov, P., Li, Q., Lin, C.-H., Chawla, G., Ostrow, K., Shiue, L., Ares, M. and Black, D. L. (2007) A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. Genes & development. **21**, 1636-1652
13 Oberstrass, F. C., Auweter, S. D., Erat, M., Hargous, Y., Henning, A., Wenter, P., Reymond, L., Amir-Ahmady, B., Pitsch, S., Black, D. L. and Allain, F. H. (2005) Structure of PTB bound to RNA: specific binding and implications for splicing regulation. Science. **309**, 2054-2057
14 Simpson, P. J., Monie, T. P., Szendroi, A., Davydova, N., Tyzack, J. K., Conte, M. R., Read, C. M., Cary, P. D., Svergun, D. I., Konarev, P. V., Curry, S. and Matthews, S. (2004) Structure and RNA interactions of the N-terminal RRM domains of PTB. Structure. **12**, 1631-1643
15 Coelho, M. B., Attig, J., Bellora, N., Konig, J., Hallegger, M., Kayikci, M., Eyras, E., Ule, J. and Smith, C. W. (2015) Nuclear matrix protein Matrin3 regulates alternative splicing and forms overlapping regulatory networks with PTB. EMBO J. **34**, 653-668
16 Izquierdo, J. M., Majos, N., Bonnal, S., Martinez, C., Castelo, R., Guigo, R., Bilbao, D. and Valcarcel, J. (2005) Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. Mol Cell. **19**, 475-484
17 Kafasla, P., Lin, H., Curry, S. and Jackson, R. J. (2011) Activation of picornaviral IRESs by PTB shows differential dependence on each PTB RNA-binding domain. RNA. **17**, 1120-1131
18 Kafasla, P., Morgner, N., Poyry, T. A., Curry, S., Robinson, C. V. and Jackson, R. J. (2009) Polypyrimidine tract binding protein stabilizes the encephalomyocarditis virus IRES structure via binding multiple sites in a unique orientation. Mol Cell. **34**, 556-568
19 Mickleburgh, I., Kafasla, P., Cherny, D., Llorian, M., Curry, S., Jackson, R. J. and Smith, C. W. (2014) The organization of RNA contacts by PTB for regulation of FAS splicing. Nucleic acids research. **42**, 8605-8620
20 Liu, H., Zhang, W., Reed, R. B., Liu, W. and Grabowski, P. J. (2002) Mutations in RRM4 uncouple the splicing repression and RNA-binding activities of polypyrimidine tract binding protein. RNA. **8**, 137-149
21 Maynard, C. M. and Hall, K. B. (2010) Interactions between PTB RRMs induce slow motions and increase RNA binding affinity. Journal of molecular biology. **397**, 260-277
22 Sharma, S., Maris, C., Allain, F. H. and Black, D. L. (2011) U1 snRNA directly interacts with polypyrimidine tract-binding protein during splicing repression. Mol Cell. **41**, 579-588

23 Hegele, A., Kamburov, A., Grossmann, A., Sourlis, C., Wowro, S., Weimann, M., Will, C. L., Pena, V., Luhrmann, R. and Stelzl, U. (2012) Dynamic protein-protein interaction wiring of the human spliceosome. Mol Cell. **45**, 567-580

24 Hein, M. Y., Hubner, N. C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I. A., Weisswange, I., Mansfeld, J., Buchholz, F., Hyman, A. A. and Mann, M. (2015) A human interactome in three quantitative dimensions organized by stoichiometries and abundances. Cell. **163**, 712-723

25 King, H. A., Cobbold, L. C., Pichon, X., Poyry, T., Wilson, L. A., Booden, H., Jukes-Jones, R., Cain, K., Lilley, K. S., Bushell, M. and Willis, A. E. (2014) Remodelling of a polypyrimidine tract-binding protein complex during apoptosis activates cellular IRESs. Cell death and differentiation. **21**, 161-171

26 Gooding, C., Edge, C., Lorenz, M., Coelho, M. B., Winters, M., Kaminski, C. F., Cherny, D., Eperon, I. C. and Smith, C. W. (2013) MBNL1 and PTB cooperate to repress splicing of Tpm1 exon 3. Nucleic acids research. **41**, 4765-4782

27 Robinson, F. and Smith, C. W. (2006) A splicing repressor domain in polypyrimidine tract-binding protein. J Biol Chem. **281**, 800-806

28 Gueroussov, S., Gonatopoulos-Pournatzis, T., Irimia, M., Raj, B., Lin, Z. Y., Gingras, A. C. and Blencowe, B. J. (2015) An alternative splicing event amplifies evolutionary differences between vertebrates. Science. **349**, 868-873

29 Huttelmaier, S., Illenberger, S., Grosheva, I., Rudiger, M., Singer, R. H. and Jockusch, B. M. (2001) Raver1, a dual compartment protein, is a ligand for PTB/hnRNPI and microfilament attachment proteins. The Journal of cell biology. **155**, 775-786

30 Gromak, N., Rideau, A., Southby, J., Scadden, A. D. J., Gooding, C., Hüttelmaier, S., Singer, R. H. and Smith, C. W. J. (2003) The PTB interacting protein raver1 regulates alpha-tropomyosin alternative splicing. The EMBO journal. **22**, 6356-6364

31 Rideau, A. P., Gooding, C., Simpson, P. J., Monie, T. P., Lorenz, M., Hüttelmaier, S., Singer, R. H., Matthews, S., Curry, S. and Smith, C. W. J. (2006) A peptide motif in Raver1 mediates splicing repression by interaction with the PTB RRM2 domain. Nature structural & molecular biology. **13**, 839-848

32 Henneberg, B., Swiniarski, S., Sabine, B. and Illenberger, S. (2010) A conserved peptide motif in Raver2 mediates its interaction with the polypyrimidine tract-binding protein. Experimental cell research. **316**, 966-979

33 Joshi, A., Coelho, M. B., Kotik-Kogan, O., Simpson, P. J., Matthews, S. J., Smith, C. W. and Curry, S. (2011) Crystallographic analysis of polypyrimidine tract-binding protein-Raver1 interactions involved in regulation of alternative splicing. Structure. **19**, 1816-1825

34 Corsini, L., Bonnal, S., Basquin, J., Hothorn, M., Scheffzek, K., Valcarcel, J. and Sattler, M. (2007) U2AF-homology motif interactions are required for alternative splicing regulation by SPF45. Nat Struct Mol Biol. **14**, 620-629

35 Kielkopf, C. L., Rodionova, N. A., Green, M. R. and Burley, S. K. (2001) A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer. Cell. **106**, 595-605

36 Lahmann, I., Fabienke, M., Henneberg, B., Pabst, O., Vauti, F., Minge, D., Illenberger, S., Jockusch, B. M., Korte, M. and Arnold, H. H. (2008) The hnRNP and cytoskeletal protein raver1 contributes to synaptic plasticity. Experimental cell research. **314**, 1048-1060
37 Castelo-Branco, P., Furger, A., Wollerton, M., Smith, C., Moreira, A. and Proudfoot, N. (2004) Polypyrimidine tract binding protein modulates efficiency of polyadenylation. Molecular and cellular biology. **24**, 4174-4183
38 Brunquell, J., Yuan, J., Erwin, A., Westerheide, S. D. and Xue, B. (2014) DBC1/CCAR2 and CCAR1 Are Largely Disordered Proteins that Have Evolved from One Common Ancestor. Biomed Res Int. **2014**, 418458
39 Close, P., East, P., Dirac-Svejstrup, A. B., Hartmann, H., Heron, M., Maslen, S., Chariot, A., Soding, J., Skehel, M. and Svejstrup, J. Q. (2012) DBIRD complex integrates alternative mRNA splicing with RNA polymerase II transcript elongation. Nature. **484**, 386-389
40 Coelho, M. B., Attig, J., Ule, J. and Smith, C. W. (2016) Matrin3: connecting gene expression with the nuclear matrix. Wiley interdisciplinary reviews. RNA
41 Guo, W., Schafer, S., Greaser, M. L., Radke, M. H., Liss, M., Govindarajan, T., Maatz, H., Schulz, H., Li, S., Parrish, A. M., Dauksaite, V., Vakeel, P., Klaassen, S., Gerull, B., Thierfelder, L., Regitz-Zagrosek, V., Hacker, T. A., Saupe, K. W., Dec, G. W., Ellinor, P. T., MacRae, C. A., Spallek, B., Fischer, R., Perrot, A., Ozcelik, C., Saar, K., Hubner, N. and Gotthardt, M. (2012) RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. Nature medicine. **18**, 766-773
42 Maatz, H., Jens, M., Liss, M., Schafer, S., Heinig, M., Kirchner, M., Adami, E., Rintisch, C., Dauksaite, V., Radke, M. H., Selbach, M., Barton, P. J., Cook, S. A., Rajewsky, N., Gotthardt, M., Landthaler, M. and Hubner, N. (2014) RNA-binding protein RBM20 represses splicing to orchestrate cardiac pre-mRNA processing. The Journal of clinical investigation. **124**, 3419-3430
43 Zeitz, M. J., Malyavantham, K. S., Seifert, B. and Berezney, R. (2009) Matrin 3: chromosomal distribution and protein interactions. Journal of cellular biochemistry. **108**, 125-133
44 Wagner, E. J. and Garcia-Blanco, M. A. (2001) Polypyrimidine tract binding protein antagonizes exon definition. Molecular and cellular biology. **21**, 3281-3288
45 Pires, D. E., Ascher, D. B. and Blundell, T. L. (2014) mCSM: predicting the effects of mutations in proteins using graph-based signatures. Bioinformatics. **30**, 335-342
46 Pires, D. E., Blundell, T. L. and Ascher, D. B. (2015) Platinum: a database of experimentally measured effects of mutations on structurally defined protein-ligand complexes. Nucleic acids research. **43**, D387-391
47 Pires, D. E., Chen, J., Blundell, T. L. and Ascher, D. B. (2016) In silico functional dissection of saturation mutagenesis: Interpreting the relationship between phenotypes and changes in protein stability, interactions and activity. Sci Rep. **6**, 19848
48 Gooding, C., Kemp, P. and Smith, C. W. (2003) A novel polypyrimidine tract-binding protein paralog expressed in smooth muscle cells. J Biol Chem. **278**, 15201-15207
49 Blatter, M., Dunin-Horkawicz, S., Grishina, I., Maris, C., Thore, S., Maier, T., Bindereif, A., Bujnicki, J. M. and Allain, F. H. (2015) The Signature

of the Five-Stranded vRRM Fold Defined by Functional, Structural and Computational Analysis of the hnRNP L Protein. Journal of molecular biology. **427**, 3001-3022
50 Salton, M., Elkon, R., Borodina, T., Davydov, A., Yaspo, M. L., Halperin, E. and Shiloh, Y. (2011) Matrin 3 binds and stabilizes mRNA. PloS one. **6**, e23882
51 Wollerton, M. C., Gooding, C., Wagner, E. J., Garcia-Blanco, M. A. and Smith, C. W. (2004) Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. Mol Cell. **13**, 91-100

Figure legends

Figure 1. A post-transcriptional network controlling expression of PTBP1, 2 and 3. Full length PTBP1 protein causes skipping of its own exon 11 (top), the equivalent exon 10 in PTBP2 (lower left) and exon 2 of PTBP3 (lower right). Skipping of PTBP1 exon 11 and PTBP2 exon 10 is frame-shifting, leading to insertion of a premature termination codon (stop sign) and Nonsense Mediated Decay. In most cell types, the cellular levels of PTBP1 cause a small amount of PTBP1 exon 11 skipping, limiting its overexpression [51], but almost complete skipping of PTBP2 exon 10, effectively switching PTBP2 expression off. PTBP1 also causes almost complete skipping of PTBP3 exon 2, which leads to production of two N-terminally truncated isoforms both lacking the N-terminal nuclear localization signal (red bar), and one of which lacks most of the first RRM domain. The relative amounts of the competing pathways in the presence of normal levels of PTBP1 are indicated by the thickness of the diagonal dashed lines. This network can be modulated by post-transcriptional down-regulation of PTBP1 [7, 11].

Figure 2. PTBP1 domain organization and protein interactions

A) Schematic of PTBP1 domain organization, with four RNA Recognition Motif (RRM) domains, and N-terminal region containing a bipartite nuclearl localization signal and nuclear export signal (red), and a 26 amino acid alternatively spliced segment (grey) between RRMs 2 and 3. Lower: interaction of the dorsal surface of RRM2 with the peptide SLLGAPP [33]. Peptide: cyan, RRM2 green, Y247 grey.

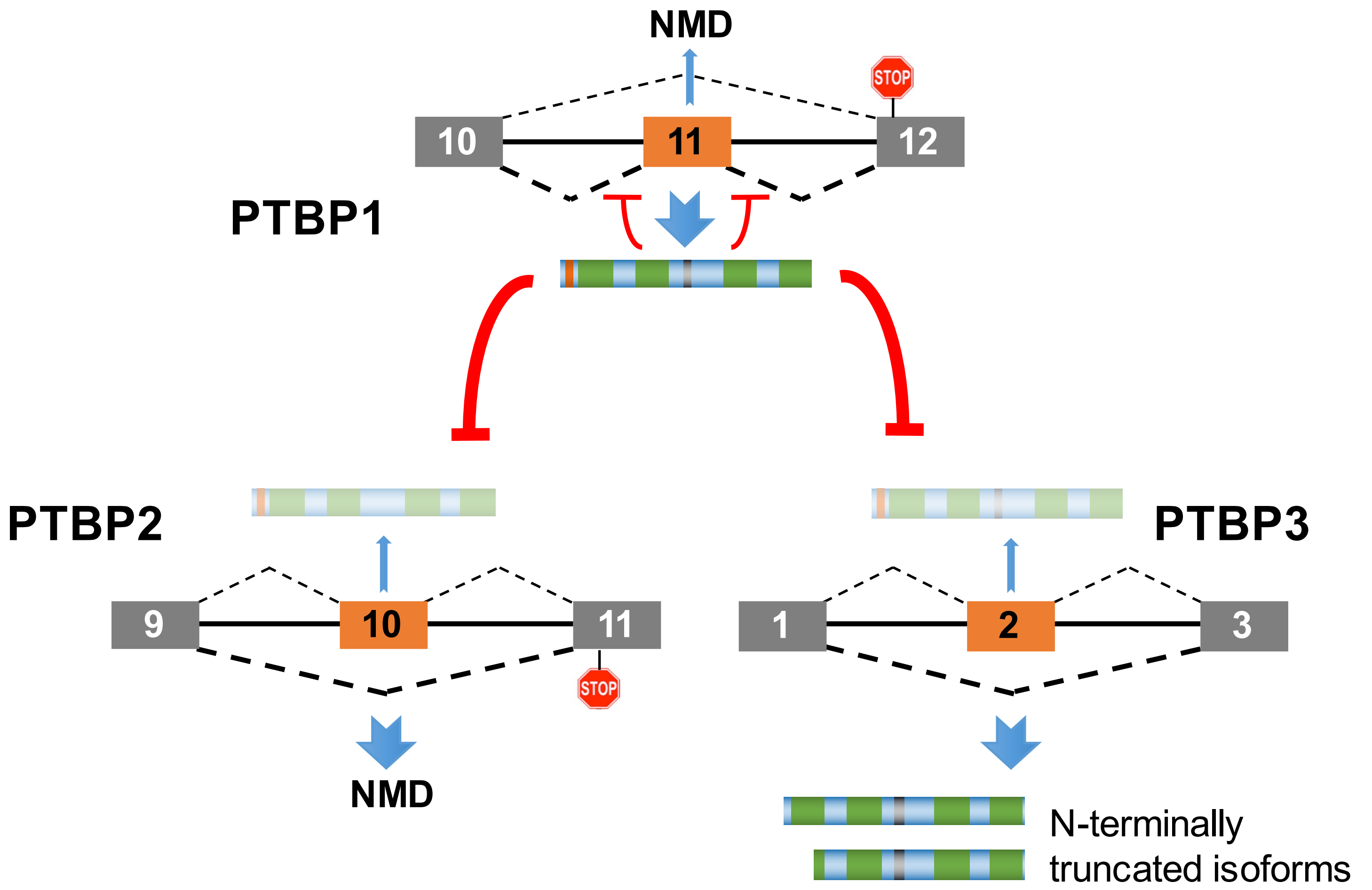
B) Proteins with potential PRI motifs identified to interact with PTBP1 RRM2 in a Y247Q sensitive manner [15]. Mismatches from the current PRI consensus are indicated by lower case. The Raver1 PRI motif in parentheses is necessary for full Raver1 activity (C. Gooding, unpublished observations), but has not been shown to interact with PTBP1. SUGP2 is included on the basis of mCSM-PPI predictions (Fig 3A).

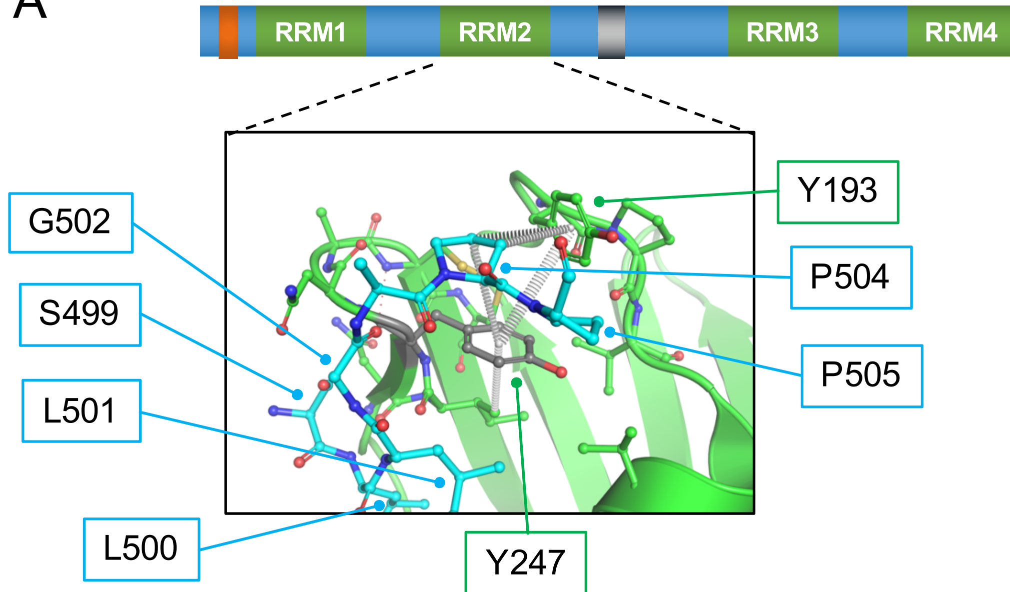
C) GST-PTBP1 pull-down of [^{35}S] methionine labelled *in vitro* translated PRI peptides. Upper row, input; middle row, GST-PTBP1 pulldown; lower row, GST pulldown.

Figure 3. Computational mutagenesis of PRI-RRM interface

A) and B) Heatmaps of mCSM-PPI predictions of effects on protein-peptide affinity of mutations in VSLLGAPP peptide (A) and PTBP1 RRM2 (B) mutations. Red, destabilizing; blue, stabilizing.

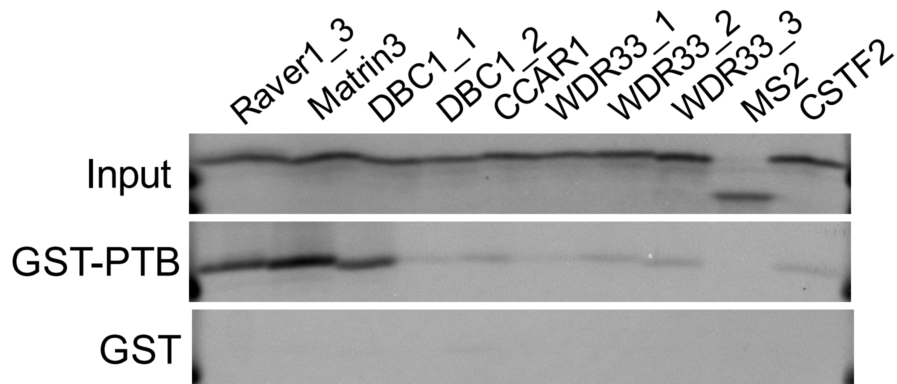
C) Alignment of critical PRI-interacting residues of PTBP1 with equivalent residues in PTBP2, PTBP3, hnRNPL and hnRNPLL. Cells are colour-coded to indicate whether amino acid changes in hnRNPL and LL were predicted to be stabilizing or destabilizing in the PTBP1-PRI interaction in panel B. Numbers in top row are amino acids positions in human PTBP1.

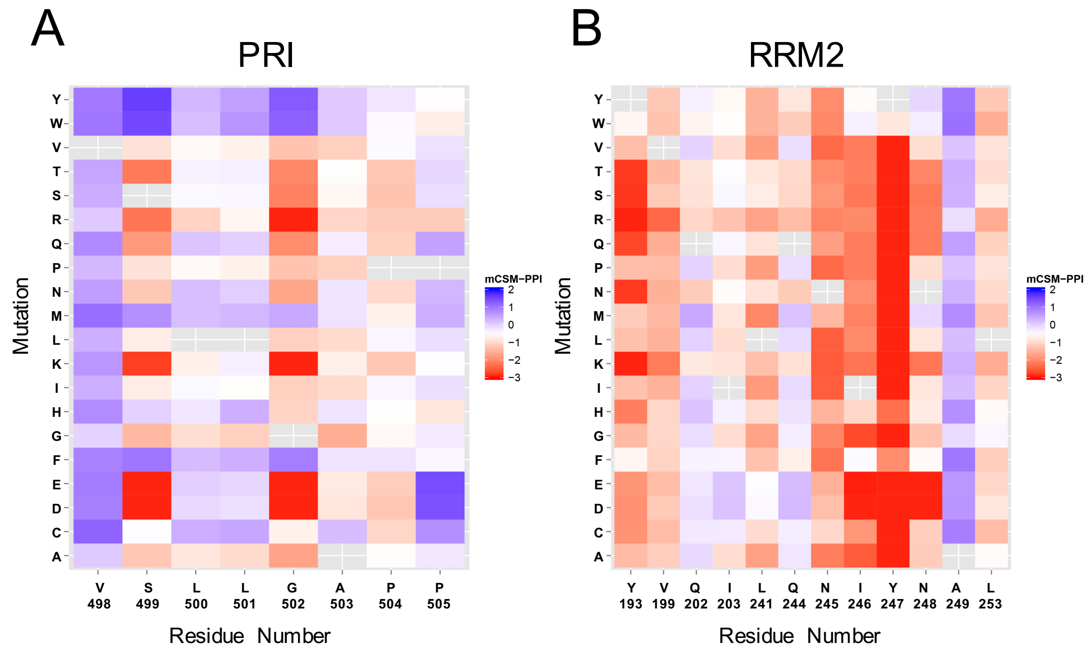

NMD
STOP
10
11
12
PTBP1
PTBP2
PTBP3
9
10
11
STOP
NMD
1
2
3
N-terminally
truncated isoforms



B

PTBP1 Interacting Protein	PRI matches	Known Functions
Raver1	GLLGAPP, GILGDsP, (pLLGELP) SLLGEPP, GLLGLsP	Splicing regulator?
Matrin3	GILGPPP	Splicing regulator, RNA export
DBC1/CCAR2	SLLGPPP, GILGAqP	Splicing regulator, transcription & chromatin regulator
CCAR1	SLLGAsP	Spliceosome component, transcription & chromatin regulator
WDR33	pLLGPqP, GLLGHgP pLLGDgP	3' end processing/polyadenylation
CSTF2	GLLGDAP	3' end processing/polyadenylation
SUGP2	SLLGKge	RS family splicing factor





C

	193	199	202	203	241	244	245	246	247	248	249	253
PTBP1	Y	V	Q	I	L	Q	N	I	Y	N	A	L
PTBP2	Y	V	Q	I	L	Q	N	I	Y	N	A	L
PTBP3	Y	V	Q	I	L	Q	N	I	Y	N	A	L
hnRNPL	Y	V	T	I	L	A	D	I	Y	S	G	L
hnRNPLL	Y	V	T	V	L	A	D	I	Y	A	G	L